DOCUMENT RESUME

ED 206 707 TM 810 609

AUTHOR Caulley, Darrel N., Ed.
TITLE Problem Case Descriptions. Research on Evaluation Program: Paper and Report Series.
INSTITUTION Northwest Regional Educational Lab., Portland, Oreg.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
REPORT NO NWREL-PR-49
PUB DATE Jan 81
CONTRACT 400-80-0105
NOTE 69p.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *Accountability; *Case Studies; Data Collection; *Educational Assessment; Elementary Secondary Education; *Evaluation Methods; Evaluators; Financial Support; Program Evaluation; School District Autonomy; *State Departments of Education
IDENTIFIERS *Evaluation Problems; Evaluation Utilization

ABSTRACT

One of a series of reports concerned with the creation of new evaluation methodologies for use in education, this report contains a collection of thirteen brief statements of problems encountered first-hand by evaluation practitioners in state departments of education (SEAs). Its intent is to make available practitioner statements so that further developmental work may be grounded in first-hand accounts. Practitioners were requested to include, where appropriate, material concerning: methodological problems and methods requiring improvement; proposals for new methods, and any problems involved in the development of such; materials and training needed to improve methods; personnel, management, planning, resource and interagency communication problems. The case descriptions are varied both in style and the problem cases they divulge. Included are such wide-ranging topics as: SEA and school district autonomy; obtaining and allocating funds for evaluation purposes; the evaluation needs of different audiences; data collection problems and quality control; the maximization of evaluation utilization. (Author/AEF)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

X This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

paper and report series

Research on Evaluation Program

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

M. M. Rogers

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Northwest Regional Educational Laboratory
710 S.W. Second Avenue
Portland, Oregon 97204
Telephone: (503) 248-6800

No. 49 PROBLEM CASE DESCRIPTIONS

DARREL N. CAULLEY, Editor

January 1981

Nick L. Smith, Director
Research on Evaluation Program
Northwest Regional Educational Laboratory
710 S.W. Second Avenue, Portland, Oregon 97204

Published by the Northwest Regional Educational Laboratory, a private nonprofit corporation. The work upon which this publication is based was performed pursuant to Contract No. 400-80-0105 of the National Institute of Education. It does not, however, necessarily reflect the views of that agency.

PREFACE

The Research on Evaluation Program is a Northwest Regional Educational Laboratory project of research, development, testing, and training designed to create new evaluation methodologies for use in education. This document is one of a series of papers and reports produced by program staff, visiting scholars, adjunct scholars, and project collaborators--all members of a cooperative network of colleagues working on the development of new methodologies.

What kinds of problems do evaluators encounter in the course of their work? This report contains a collection of thirteen brief statements (2-9 pages) of problems encountered by evaluation practitioners in state departments of education. These problem case descriptions, prepared by practitioners themselves, provide insight into the difficulties of state level evaluation practice.

Nick L. Smith, Editor
Paper and Report Series

CONTENTS

		Page
Introduction		1
Problem Case Descriptions		4
1.	Nonpublic Student Auxiliary Services Program	4
2.	Evaluation in the State of Steiner	13
3.	The Evaluation Needs of Differing Audiences	15
4.	The State Agency and Third Party Status	21
5.	The Problem With Positives	24
6.	Cooperation From Schools in Collecting Data	28
7.	The Politics of Evaluation: A Bilingual Case Study	32
8.	Quality Control	39
9.	Basic Skills Evaluation	42
10.	Monitoring - A Threat to Evaluation?	48
11.	Betwixt and Between	52
12.	Maximizing Assessment Results Utilization	56
13.	Coleman Assessment Program	60

PROBLEM CASE DESCRIPTIONS

Introduction

What kinds of problems do evaluators encounter in the course of their work? This report contains a collection of thirteen such problems as reported by evaluation practitioners in state education agencies (SEAs). These cases were assembled by Dr. Darrel N. Caulley of the Research on Evaluation Program so that program staff would have a better understanding of the difficulties encountered by state department evaluators. Assumptions about the nature of evaluation practice, its problems and constraints, underlie much of the work of the program. These collected practitioner statements enable program staff to test the validity of staff views of practice and to ground further work in first-hand accounts of evaluation practice.

A number of evaluators in state departments of education were paid a nominal fee to prepare brief (2-9 pages) statements of a problem they had encountered in their work. They were asked to describe an actual problem in sufficient detail so that others could understand the nature of the problem, its context and implications. The writers were allowed to select any problem they wished and to present it in the form they thought most useful. (Of course, names, dates, etc. were altered to insure anonymity.) To aid their efforts, writers were sent sample problem case descriptions obtained from business education materials. The following list of possible problem topics was also provided:

a. What methodological problems do SEAs currently have?

b. What new methods are needed?

c. Which methods need to be improved?

d. What problems are there in the development of new methods?

e. What problems are there in the implementation of new methods?

f. What kinds of materials and training systems are needed to help SEAs improve methods?

g. What serious personnel/staffing problems does your SEA evaluation unit have?

h. What serious management problems does your SEA evaluation unit have?

i. What serious planning problems does your SEA evaluation unit have?

j. What important resources problems does your SEA evaluation unit have?

k. What important reporting problems does your SEA evaluation unit have?

l. What important interagency communication problems does your SEA evaluation unit have?

The thirteen problem statements which follow vary considerably in topic, length, and format. We are grateful to the following individuals who provided these problem case descriptions:

Mary Ann Awad	New York Department of Education
Bill Burson	California Department of Education
Alex Hazelton	Alaska Department of Education
Jerry Hutchinson	Mississippi Department of Education
Thomas Kerins	Illinois Office of Education
Ann Kraetzer	Colorado Department of Education
George Malo	Tennessee Department of Education
Claudia Merkel-Keller	New Jersey Department of Education
Lyn Nachman	Minnesota Department of Education
Michael Plog	Illinois Office of Education
Norman Stenzel	Illinois Office of Education
Donna Van Kirk	Washington Department of Education

Dr. Caulley originally intended to supplement this collection of case descriptions with additional analysis and commentary. Due to a serious and lengthy illness, however, he has been unable to do so. Hopefully, he will be able to return to this task in a few months. In the meantime, these problem statements have been gathered here for use by program staff.

Nick L. Smith

Problem Case Description No. 1

Nonpublic Student Auxiliary Services Program

Background

Why Mary can't read and why Johnny can't add are serious concerns facing educators at all levels--local, state and national.

A state department of education in one of the greater metropolitan areas of the nation is developing and implementing significant remediation and compensatory programs to deal with the large numbers of its students falling below the state-established minimum basic skills performance levels. By state mandate all of those students falling below the state standard must be served by a remedial program which can be funded, in full or in part, by local, state and federal funds or a combination of these. As can be expected, the state department of education implements a very sizeable Title I program which serves both public and nonpublic students. Approximately 77 million dollars are allocated to the state by the federal government every year for the provision of Title I programs and projects.

The State Legislature has also made a commitment to provide remedial programs to the state's students through the allocation of about 68 million dollars for the State Compensatory Education Program. The State Compensatory Education Program is administered by the State Department of Education and is aimed at the public sector. In 1977, the State Legislature also passed two laws with a funding package of about 9 million dollars which provided auxiliary services and general services for nonpublic students. This legislation was a significant step forward in terms of the allocation of public monies to service the nonpublic sector. At the heart of the issue is the separation of church and state.

As can be seen from the previous discussion, this state's budgetary commitment to providing compensatory services to its students is in the neighborhood of 160 million dollars. Of course, from a policy and decision making level, both state and national, the question arises, "Is the program working?"; that is to say, "Are the children learning more as a result of the program?" These accountability and evaluation questions, as well as others, gave rise to the development of the Title I Evaluation and Reporting System (TIERS) at the national level through pressure from the U.S. Congress.

The TIERS system attempts to formalize data collection and reporting across the states through the implementation of three outcome evaluation models or designs coupled with the gathering of other program and process data.

The three outcome evaluation models or designs are as follows:

Model A: the norm-referenced model,
Model B: the control-group model,
Model C: the special regression model.

All three models are designed to be used with any valid and reliable norm-referenced test or criterion-referenced test. Additionally, each of the models requires both pre-testing and post-testing and imposes some special conditions and restrictions on the testing itself. The three models each provide data on an observed post-treatment performance measure and an estimate of what that performance would have been without the program (i.e., without the treatment).

This state viewed the models proposed and then mandated in 1979 for Title I programs as a viable method to look at evaluation for all of its basic skills preventive and remedial programs in the state regardless of funding source. The state took a bold step and mandated the use of the Title I Evaluation and Reporting System (TIERS) for all of its compensatory programs operating in the state. The state has been successful in the implementation of TIERS for its Title I programs and for its State Compensatory Education in the public sector. Plans are

currently underway to augment the basic output design with other types of program and process data.

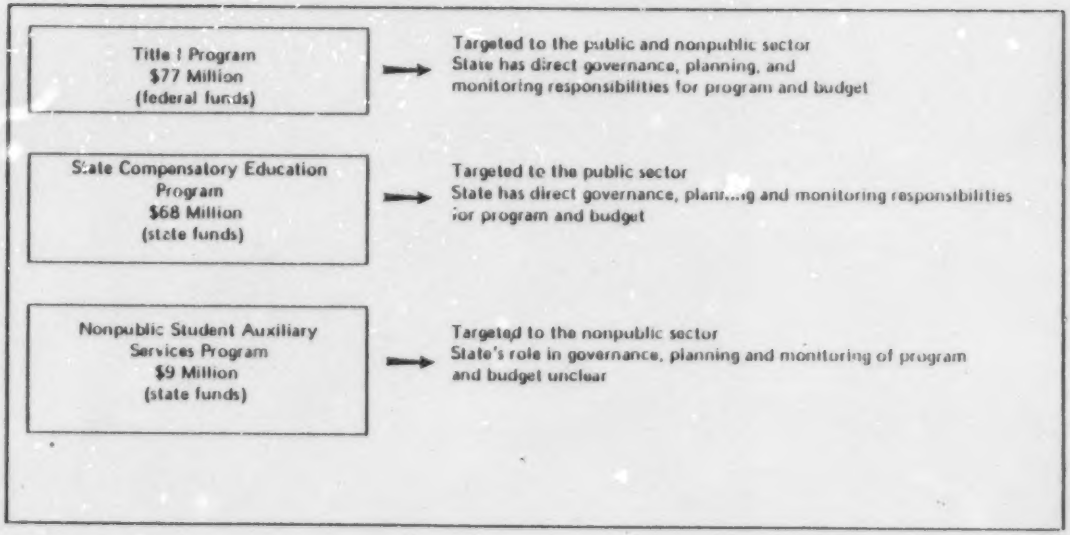
Issue

At the core of this issue, however, is the state's role in the planning, development, and implementation of an evaluation system including TIERS for the compensatory education services provided under the Nonpublic Student Auxiliary Services Program (see Figure 1).

The issue is a complex one, since it not only deals with (1) the issue of state control and governance in terms of the separation of church and state, but also with (2) philosophical concerns which suggest that if an agency accepts monies from the state that agency must be accountable to the state for those monies and must be governed by the implementation rules for the use of those monies. Added to the governance issue is (3) the issue of program planning and evaluation. At the state level, one needs to know if the program is working, and indeed, if program services are reaching the intended audiences. At present, the State Department of Education is at a loss to be able to say anything much beyond the total outlay of monies by category by school district. State monitoring of the program is virtually non-existent. Again, by default (or at leat post-hoc), program evaluation planning may give substance and shape to the overall planning, implementation, and delivery of the particular program in question.

FIGURE I

MAJOR STATE BASIC SKILLS PREVENTIVE
AND REMEDIAL PROGRAMS AND FUNDING SOURCE


Title I Program
$77 Million
(federal funds)
Targeted to the public and nonpublic sector
State has direct governance, planning, and monitoring responsibilities for program and budget
State Compensatory Education Program
$68 Million
(state funds)
Targeted to the public sector
State has direct governance, planning and monitoring responsibilities for program and budget
Nonpublic Student Auxiliary Services Program
$9 Million
(state funds)
Targeted to the nonpublic sector
State's role in governance, planning and monitoring of program and budget unclear
Evaluation Problem Area
(focus of discussion)

Description of the Nonpublic Student Auxiliary Services Program

The two state laws which comprise the funding package for services to nonpublic students can be partitioned into the following subsections:

Auxiliary services for nonpublic students

1. "Compensatory education" means preventive and remedial programs in basic communication and computational skills as set forth in the state administrative code.

2. "Supportive services for acquiring communication proficiency in the English language for children of limited English-speaking ability" means programs in English as a second language.

3. "Supplementary instruction" means instruction provided for a pupil classified pursuant to state law as handicapped; it is given in addition to the regular instructional program of such a pupil, as set forth in the state administrative code.

4. "Home instruction" means individual instruction given in lieu of regular classroom instruction to a pupil who is unable to attend school because of illness or injury, as set forth in the state administrative code.

5. Examination and classification of potentially handicapped pupils (i.e., child study team services).

General services for nonpublic students

6. Corrective speech services (articulation disorders).

In order to focus the discussion, only evaluation problems dealing with Category 1 (compensatory education) will be dealt with in the subsequent discussion.

Services falling under "compensatory" correspond to the generally accepted definition of a preventive and remedial basic skills programs as provided under the state law. Therefore, only services which are "compensatory" would be subsumed under the program evaluation procedures (TIERS) currently being implemented statewide.

The one major problem with program evaluation under the state law governing nonpublic education is the lack of reference to program evaluation requirements in either the law itself or in the interpretive and guideline materials prepared by the State Department:

Problem with lack of reference to program evaluation in law

"At the close of the school year, the district board of education shall submit to the Commissioner a report describing the classification and corrective services provided by the district board of education purusant to state law. The report shall be completed in a manner prescribed by the Commissioner and shall include, but not be limited to, such information as the classification and corrective service provided, numbers of nonpublic school pupils served, frequency and/or amount of the service, and facilities utilized."

There are several problems or concerns raised by provisions in the state law governing nonpublic education concerning the implementation and management of services:

* Services must be provided in a non-sectarian facility (i.e., students much receive services away from their usual environment).

* Services must be arranged for and managed by the public school and may not include use of any staff employed by the nonpublic facility.

Problems with implementation and management of services

Services may be arranged either through contracting, hiring of staff by the public school, or through cooperatives among more than one public school district.

". . . Services may be delivered in a variety of ways, depending upon such factors as number of pupils, kinds of services, location of facilities, personnel available, logistics, funds available, etc. Some of these ways include the following:

* Districts themselves providing services to all eligible pupils for whom these districts are responsible;

* Two or more districts cooperating to provide services to all eligible pupils attending nonpublic schools located within their respective districts, whether or not the pupils actually reside in the same district where the nonpublic school they attend is located;

Delivery Strategies

* Districts providing services through a county educational services commission;

* Districts contracting with an educational improvement center to provide services; and

* Districts contracting with a non-sectarian private school to provide services"

Each of the above represents any number of variables which may have an impact on the services. With the variations that can occur, assumptions regarding consistency in treatment conditions fall apart. Measurement of program impact could be aggregated

only if reasonable controls are built in to retrieve relevant information on setting, types of services, etc. Given the fact that services must occur outside of the nonpublic school, the potential for a breakdown in communication between the regular teachers and the compensatory staff increases.

Counter to the provisions under the state law governing nonpublic education are the provisions for nonpublic services under P.L. 89-10 (as amended). The ESEA Title I program, which requiring management by the public school, fosters much more consistent services in that:

1. Services are provided on-site in the nonpublic facility.

2. Services are provided by staff specifically identified as instructional personnel for the nonpublic facilities (since under 192, these personnel are not employees of the nonpublic school). Services are provided through coordination with the administrators and staff in the nonpublic school.

The Title I program also represents a joint planning effort between public and nonpublic schools in that a single program plan is developed, needs assessment is coordinated, and program evaluation procedures are designed for both the public and nonpublic components (while these may differ in terms of specifics, provision for evaluation is present for both).

Another general problem with the state law governing nonpublic education is limited funds available. By the time suitable facilities are formed, transportation is arranged, etc., the number of students who could be effectively served may be very small in some cases. The nonpublic monies are used to pay all of these costs. Controls on the size, scope, and quality of services appear to be more limited than under Title I.

The fact that services must occur "off-site" precludes any provision for comparison groups. This means that the only reasonable model that might be appropriate for evaluating compensatory services under the nonpublic funding category would

be TIERS A1 (TIERS A2 would not be appropriate unless a large N were being served and three test administrators could be scheduled). If TIERS A1 is required, then issues pertaining to conflicts with on-going testing in either the public or nonpublic schools arise. The question or consistency of services at various sites, etc., all enter into the picture.

Evaluation Questions, Evaluation Problem Areas, and Evaluation Needs

1. Re-write (impact) on legislation to provide for a clear evaluation mandate for those programs governed under the state law for nonpublic education.

2. Design a comprehensive state evaluation plan for the Nonpublic Student Auxiliary Services Program including:
   * How are the appropriated funds spent by each categorical area?
   * How much instructional time does each student receive in each area?
   * What are the most frequently used models for the delivery of these services?
   * Which models are the most cost-effective in terms of their operations delivery of services and in terms of student impact?

3. Develop a concept paper on the state governance role over the nonpublic sector.

4. Refine the Title I Evaluation and Reporting System (TIERS) to accomodate those special needs and problem areas defined in the previous section.

5. Develop a comprehensive program and budget evaluation, reporting, auditing, and monitoring system.

6. Develop an evaluation training plan for staff in the nonpublic sector.

Problem Case Description No. 2
Evaluation in the State of Steiner*

Within Steiner State Department of Education, evaluation comes under the Office of Planning and Evaluation. According to the coordinator of the office, the office is fortunate to be staffed with highly qualified personnel. The office staff is made up of three individuals, each with a terminal degree in the field of education, and two highly skilled secretaries. According to the coordinator, methodological problems that might be prevalent among other state education agencies do not seem to be a concern in Steiner. The coordinator feels that the expertise of the staff within the Office of Planning and Evaluation negates methodological problem areas.

The biggest problem area confronting the Office of Planning and Evaluation is the problem of acquiring appropriate funding sources to effect proposed studies. There have been many examples where study proposals have not been funded.

Management within the Steiner State Department of Education has been very supportive of the Office of Planning and Evaluation. The typical routine followed by this office is to prepare a formal proposal, obtain proper authorization before proceeding, and seek an appropriate funding source. Almost always the appropriate funding source becomes the obstacle that hinders implementation. Attached is a brief synopsis of the funding problem prepared by the coordinator for one of the state legislators.

The Problem

Compared with other states, Steiner seems to be the recipient of a disproportionate share of federally allocated funds which are administered through educational projects sponsored by the National Institute of Education (NIE) through research grants.

The Cause

Projects funded through NIE usually result from written proposals which emerge successfully from a screening process. Policies of NIE stipulate competition among potential recipients for the available funds. The very fact that competitiveness exists works disfavorably for Steiner. For example, some state education agencies have on their staffs trained proposal writers. In Steiner, the responsibility for writing a proposal

---

*A fictitious name for a state.

is usually assigned to a person in the office who most nearly relates to the topic under consideration. That person presumably already has full-time responsibilities and will likely give only token effort to the task of writing a proposal, especially when he realizes that should the proposal be funded he will be given the responsibility of administering the project in addition to his regular duties. Obviously, this procedure lessens the prospects of obtaining an ultimately funded proposal which has been subjected to the rigors of competition. If Steiner could afford the luxury of employing a proposal writer, the salary that that individual could expect would be far below the salaries of comparable positions in other states. Again, competition would likely rule out the possibility of a project being funded for Steiner. Some exceptions do exist. However, when one compares the total amount of money received by states through educational research grants, he will find that Steiner does not compare favorably with those states that have more wealth and more skilled manpower for developing technical proposals.

The Cure

If proposed projects appear necessary or useful for disbursing funds, the money could be allocated to states on the basis of formulas which address such factors as school enrollment, population, per capita income, etc. Obviously, this procedure would negate the necessity of competition for funds for educational research projects. This does not mean, however, that competition for grants should be totally eliminated. Certainly, by the nature of some proposed projects competition is desirable. Discretion is necessary.

Problem Case Description #3

The Evaluation Needs of Differing Audiences

Two and a half years ago, a state department of education instituted a full-time position for an internal program evaluation consultant. This position was designed as an alternative to hiring independent (i.e., external) contractors to perform evaluations required as part of federally or state funded projects.

Since then, the demand for these services has more than doubled. The evaluation staff is now grappling with issues related to increasing its effectiveness within the organization.

In contrast with other state education agencies, this Department is comparatively small and has limited regulatory responsibilities. It is comprised of approximately 125 professional staff who serve the State Board of Education. Its major function is to provide leadership to the state's public school system.

The State Board of Education has five members who are the elected representatives of the state's five congressional districts. The Board's responsibilities are to distribute state and federally apportioned funds to the schools, submit recommendations on education improvements to the Governor and General Assembly, and to appraise the work of the Commissioner of Education (whom the Board appoints), the Department of Education, and the state's public school system.

Because the state strongly adheres to principles of local control and autonomy, the Department staff's primary functions are to provide leadership and technical assistance to school districts and to administer federal and state categorical educational programs. These services are organized in four offices of the Department, each headed by an assistant commissioner, and the office of the Commissioner. The Commissioner of Education and Assistant Commissioners form the Department's Executive Committee. Within the five offices are

thirty units, each specialized to provide either program services to the schools or support services to the state Department of Education.

The Planning and Evaluation Unit has the prime responsibility to assist the Commissioner in preparation of the Department's budget. A modified Program Planning and Budget System (PPBS) called Planning and Management System (PAMS) was developed by the Department to coordinate the budget process. PAMS includes a self-evaluation component for each program operated by the Department.

This self-evaluation component is geared toward the information needs of the state legislature and does not provide the level of detail needed by managers of categorical projects.

Prior to the availability of internal evaluation consultants, evaluation needs were met in two ways. Program evaluations were conducted by the project staff or through contracts with independent consultants. In the first situation, staff typically lacked the expertise and the time to conduct more than perfunctory program reviews. Contracts with outside consultants, however, posed additional problems. The Executive Committee became disillusioned with independent contractors because their results too frequently were characterized by one or more of the following concerns:

- Lack of timeliness;
- Lack of formative evaluation;
- Overly biased in a positive direction;
- Overly technical
- Not responsive to the informational needs of the Executive Committee and State Board; and
- Lack of staff commitment to the evaluation results.

To remedy these concerns, the Department agreed to expand the evaluation role of the Planning and Evaluation Unit by adding staff with specialized evaluation training. Funding for the evaluation staff did not follow Scriven's prescription that evaluation funding not come through the program budget. Rather, the line-item entry of evaluation in federally and state funded program budgets made internal evaluations possible.

Funds earmarked for evaluation are noted by the Planning and Evaluation staff during the internal proposal review process. Project managers are contacted and informed of the evaluation services available through the Unit. If the project manager is interested, the evaluation staff initiates a process of defining evaluation objectives. This process leads to a formal contract for services between the Planning and Evaluation Unit and the project manager.

The process, which is outlined in Figure A, is designed to facilitate responsiveness to each level of management in the State Department of Education and to overcome the concerns listed above.

That the process is successful is evidenced by increased demand for the service. The process has met its objective to increase the involvement of the Executive Committee in defining evaluation objectives. All parties to the evaluation concur that benefits come from the accessibility, communication, and common understandings that the internal environment affords.

Because the evaluations are conducted by management service staff, not program staff, the objectivity of the evaluators is enhanced. The evaluation staff is directly accountable to the Planning and Evaluation Unit director who acts as a buffer and mediator to protect the evaluator's integrity. Proximity to project staff fosters close working relationships that increase the evaluator's understanding of the project staff's view of their program's purposes, goals, and problems.

However, the evaluation staff frequently feels torn between the needs of project managers and higher level managers in the Department.

Because funds available for any given study generally are limited, and the studies must meet federal or state program requirements for evaluations, the scope and depth of the studies has necessarily been restricted. While program managers tend to give high priority to formative evaluation objectives, the State Board, the Commissioner, and the Assistant Commissioners have

greater need for answers to summative evaluation questions. Federal and state requirements typically encompass both categories of evaluation objectives but funds are not adequate to provide thorough responses to either category. The evaluation staff efforts become fragmented in trying to respond to all of the information demands.

Therefore, the Planning and Evaluation Unit is seeking new ways to increase the efficiency and effectiveness of the evaluation process. Some alternatives being considered are:

FIGURE A

Evaluation Process
Planning and Evaluation

The following information represents the basic steps to be followed by consultants in the Planning and Evaluation Unit in conducting evaluation studies:

I. Establishment of Evaluation Objectives - Cumulative Steps (In writing)

   A. Proposal or Plan Review - Series of Evaluation Objectives (as appropriate)

   B. Project Manager - Verification and Development of Additional Objectives (as appropriate)

   C. Unit Director - Verification and Development of Additional Objectives (as appropriate)

   D. Commissioner - Verification and Development of Additional Objectives (as appropriate)

II. Development of Evaluation Design and Timeline

III. Development of a Contract Agreement with Project Manager

IV. Implementation of the Evaluation Study

V. Conduct a Monthly Evaluation Program Review with Project Manager and Unit Director

VI. Development of a Draft Report

VII. Development of a Final Report

VIII. Development and Presentation of an abstract of the Final Report to:

   A. Executive Committee

   B. State Board of Education

Verification & Approval by Decision Maker Guidelines

- Soliciting independent funding for studies of broader issues raised by the State Board, the Commissioner, and the Executive Committee;

- Increasing the technical assistance given by the evaluators to program staff so that program staff would assume greater responsibility for conducting formative components of the evaluations;

- Prioritizing requrests for different kinds of evaluation information;

- Increasing the use of clerical and student intern assistance for routine evaluation procedures; and

- Increasing coordination of evaluation data requirements with the Accounting Unit and the School Finance and Data Services information systems.

None of these alternatives would be easy to accomplish because fund cuts are occurring at all levels of education and present staff workloads are already overburdened.

Questions

1. What other alternatives could the Unit take to improve the responsiveness of the evaluation studies?

2. What communication systems are necessary to mediate the differing information needs of the organization's hierarchy?

3. What are the major advantages of

   a) internal evaluations?

   b) independent evaluations?

4. What are the appropriate kinds of objectives for evaluations done by

   a) project staff?

   b) management staff?

   c) independent contractors?

5. What actions should the Unit take? Why?

Problem Case Description #4
The State Agency and Third-Party Status

There are times when State Education Agency (SEA) evaluators have difficulty keeping their third-party status. George, a fictictious name for a person who was once my boss's boss, asked me to evaluate a program that he had helped institute. He had an active interest in the success of the program. The results of my evaluation were negative; the program was canceled. My report, naturally, was not the only reason for the cancellation.

For this episode, the program being evaluated is not greatly important, but it should be briefly described. The program was established to create cooperation between the SEA and some of the state universities. Personnel from the SEA were given release time (at full pay) to attend classes, work toward advanced degrees, and provide services to the universities (i.e., a graduate assistant). The universities were to make resources--professional time--available to the SEA for no charge. The professors were given release time to provide these free services for the SEA.

The major reason for the lack of success of the program relates to the use of the university resources by managers within the SEA. Most agency managers knew of the program, but did not know how to go about obtaining the free services offered. (This is a gross over-simplification, but adequate for our purposes here.)

It is important now to understand the organization of the evaluation unit within the SEA. I was an evaluator, reporting to the director or evaluation. There were other units with similar functions in the department. George was the administrator of the department. Because of this organizational structure of the SEA, I was reporting to, and (in a sense) evaluating the same person.

That was a very unusual situation in the SEA. Normally, our evaluations were conducted on federal programs operated through the SEA, such as Title I, Title IV, Special Education, Handicapped, etc. George, and other people in his department,

were not involved in the administration of any of the programs we normally evaluated. Since George reported to the Chief State School Officer, our unit (and entire department) was similar to a staff position as opposed to a line position. We were in essence third-party, disinterested evaluators housed within the SEA. To evaluate a program that George partially administered was, as stated before, very unusual. The evaluation was a special case, for a short time period, and for one report only.

Even though this evaluation was a special case, I still had difficulty keeping a third-party mental frame of reference. The difficulty, I must hasten to point out, was caused not by George, but by me. As far as I know, he had no problem with my status.

The major reason for my concern was my newness in the position I held at the time. I did not know George very well; we had then recently started working together. I was fairly new with the agency also. My lack of experience and knowledge about George and the agency caused some insecurity on my part. George did not seem to be a vindictive person, but I did not know how he would react to a potentially threatening evaluation report. As it turned out, he was indeed not vindictive, and reacted very well.

In addition to my personal difficulties, there can be some professional problems with this type of evaluation. An evaluation lacking third-party status is a no-win situation. If the final report is positive, critics of the program can easily claim the evaluator is trying to hide something. The report may have little credibility even with supporters of the program being evaluated. If the final report is negative, the evaluator may be in a less comfortable position than I was. Superiors can be threatened to a greater extent than George appeared to be. There could be charges of disloyality, or trying to do a "hatchet job" from within. Future reports and findings may not be accepted well, because of the past associations between evaluator and organizational administrator. All this reflects on the entire evaluation unit, not just the single evaluator within the unit.

The credibility of the unit can suffer, which hampers other work that is truly third-party.

The results of this evaluation were not nearly as bleak as they could have been. George recommended appropriate action be taken with regard to the program. There was no animosity between us after the report was completed.

Problem Case Description #5
The Problem With Positives

It doesn't happen often, but a few evaluation reports are entirely positive; there are no negative findings. The program smells like a rose, with no thorns attached. Non-evaluators may have difficulty imagining the questions that go through our minds in such a situation. What did I miss? Have I been co-opted by the program's ideals or personnel? Did I--intentionally or not--do a whitewash on the program? Should I do one more chi square, one more interview?

It is possible for programs to meet all the standards agreed on before the evaluation. While it may be true that other standards could be applied (or the standards themselves could be evaluated), that may involve adding new rules to the game. Also, some clients want specific information. Even then, dear Abby, why do I have such a difficult time living with no negatives?

Evaluators are tempted to search for negative results. This is a tendency we share with auditors and other third-party investigators. I would like to discuss some reasons why we are tempted to find negative results about the programs we evaluate.

First, finding negative aspects about programs validates our existence and service. We are especially validated if we find something that everyone else missed. Without our service, how would anyone know the flaws in a program? We can best see this idea of validation by defining the purpose of evaluation as assistance to decision makers. In order to "help" someone, we must start with a problem. If we have no problem, we cannot help, therefore the money spent on evaluation has been wasted. If we cannot tell a decision maker that something needs changing, we have no service to offer. Our jobs then have no reason for existence. Good grief, we could even be eliminated. After all, there are mortages to pay, children to feed, and a cat that has not hunted anything more dangerous than a paper wad.

Another reason we are tempted to search for negative findings relates to our view of our role. It is easy for an evaluator to get carried away with the image of an investigative reporter striving valiantly to uncover governmental waste and inefficiency. We have, after all, a moral obligation to the public, and especially the funding sources. Some evaluators do not view their role as similar to investigative reporters. There still remains the image of the disinterested truth seeker. We simply cannot depend on program managers to point out problems with their programs.

Finally, it is easier to write about negative findings than to write about positive findings. Positive findings tend to produce a "so what" feeling on the part of the evaluator. When we put something negative down on paper, we need a lot of ammunition. We expect disagreement from program personnel, so we have to make an extra strong case. In order to justify our findings, we must gather all the support possible from the data. This justification even includes many of those techniques we learned in graduate school, such as how to use statistical terms so no one understands what we are saying. This takes up space in our report, and we are aware that the value of a report is directly correlated with its thickness. (Well, even if we disagree, our boss believes that, so we have to turn out thick reports. Ever heard of a person getting a raise because of a thin report?)

These are some of the reasons we are tempted to search for negative results. The temptations are powerful. We feel guilty if we do not come up with at least one bad thing about a program. I think, however, we should fight the seduction of comfort in negative findings and not be ashamed to present positive findings.

Some evaluators are so ashamed of positive findings that they will present them in a reverse manner. ("There is no evidence that this program causes harm to students' reading ability.") This is almost a double negative. There is a difference between

saying we found "nothing wrong" and saying we found "something right".

Positive findings should be presented with the same intensity and fervor as negative findings. We still must justify our findings, and gather all possible support from the data. If evaluation findings (either positive or negative) are to be accepted, the rigor of any study must be evident to the readers.

Evaluators should realize that it may be helpful to point out the positive features of a program. Perhaps the program managers only get a feeling of security, not specific suggestions of things to correct. Even so, other people are involved with the program and interested in the results. Evaluation reports are read by people in funding sources, oversight groups, governing or advisory committees, as well as the program managers' superiors. While we often moan and complain about the lack of response to our reports, we really do not know how widely our words are read. Let us accept for a moment that in some cases the system works, and decision makers do indeed read our reports. A positive report can be a help to decision makers at many levels. An easy example to show this help is the case of two programs competing for the same funds.

It is important for evaluators to understand that we are not trying to sell newspapers to a jaded public. We may borrow techniques from the field of investigative reporting; we may even borrow a degree of skepticism from that field. Our purposes, however, are different. We are supposed to be discovering the nature and worth of a program. We may borrow techniques from other fields, such as an adversial court situation, but again our purposes are different. Idealism may be very helpful in some situations, but nympholepsy is not much value to an evaluator.

(I will at this point seek forgiveness from all supporters of the union between investigative reporting and evaluation. My words above show an admittedly less than perfect realization of the benefits of such a union. I am personally <u>not</u> opposed to discussions of similarities between the two fields.)

This is not a call for evaluators to search for something positive to say. It is definitely not a call to try to collect data that will show a program in a positive light. I am simply stating that the evaluator should not avoid presenting positive findings. Nor should an evaluator feel any guilt or shame because of a positive report. I started this document with the comment that entirely positive reports do not happen often. Perhaps that is as it should be. Perhaps other evaluators, like me, have searched for something negative. What about you?

Problem Case Description #6

Cooperation from Schools in Collecting Data

A State Department of Education Testing and Evaluation Unit committed itself to participate in a National Longitudinal Study of high school sophomores and seniors. Since the National Study had insufficient schools in its study for the state to generalize about students in the state, the State Department's Testing and Evaluation Unit decided to compliment the national study of 12 schools by adding an additional 50 schools--thus creating a total state sample of 62 schools.

The study design called for requesting a sophomore and senior class roster from each of the selected schools, randomly choosing 36 students from each of these class rosters, and inviting these particular students to participate in a three-hour survey. Each student who participated would be asked to complete three booklets: Identification pages; a questionnaire; an achievement test. The 50 schools chosen represented a variety of geographical locations, and a mixture of urban, suburban, and rural communities. Because of the small number of schools in the study, it was particularly important that the mix of schools be maintained in order to make some generalizations about students in the state.

To insure school cooperation with the study, an agreement of co-sponsorship was established between the Evaluation Unit and the State's Association of School Principals. This arrangement was intended to provide principals with some additional information about the study and additional incentive to cooperate. Some phone calls were made by association members to principals of selected schools to informally encourage their support.

To prepare for contacting school districts and school principals, a two-day training of evaluation staff was conducted by a member of the national office responsible for the study. During the training, staff reviewed forms, procedures, survey

booklets, etc., which the national staff prepared, and some thought was given as to how particular materials might be adjusted to suit the needs of the state's study. At the conclusion of the training, five two-person teams were formed and school assignments were made. Teams were told that a letter had already gone to district superintendents, with a carbon of the letter addressed to principals, and a letter addressed directly to principals was scheduled to be mailed the following week. In addition, teams were informed that they would receive a memo covering revised instructions of copies of revised materials within a few weeks, at which time it would be appropriate for them to begin phoning principals to 1) establish the principal's willingness to cooperate; 2) reitterate the nature of the study; 3) receive the name of the school staff contact person who would handle details. Lastly, teams were told that the school surveys needed to be concluded by May 1, which gave them a total of three months for the effort.

During the month following training, several staff spent their time reviewing all the national materials, editing where necessary, making decisions concerning changes in directions, and deciding how many copies of materials would be printed. Additional staff time was spent developing a management information system so that the entire process could be monitored.

Throughout the first month of the three-month study, certain assumptions were made about school district participation. Since letters explaining the study were mailed to district superintendents on January 30th and school principals on February 7th, no response from them was interpreted as a positive response. Believing that adequate time had elapsed for school districts to send negative responses, no efforts were made to confirm district approval other than some contacts made by the Director of Evaluation to confirm the willingness of the staff of the largest school district in the state to have seven of its twelve schools surveyed. The Director called the Assistant Superintendent and was informed that the district office required

completion of a form "Application to Conduct Research and Experimental Studies in the Brisbane* Public Schools", and so this form was completed and promptly returned. During the third week of February, the Director visited the school district office and received a verbal assurance by the Assistant Superintendent that the study would be approved. Five days elapsed and a written response arrived indicating that the district would not approve the study. This letter, received 28 days after the initial letter was sent to the District Superintendent explaining the study, came as a complete surprise. By February 28, all materials were printed, 52 boxes containing three booklets for each of the 3,[illegible] students had arrived from the national office, and teams were preparing to schedule school visits. The magnitude of the problem of having the largest school district in the state withdraw from the study was summed up quickly by the Director--the loss of that district would end the study.

During the next several days, steps were taken by the Evaluation Unit Director to open up other possibilities to salvage the study:

1. The State Superintendent of Education assured the Director that he would be willing to write a letter to the District Superintendent asking for a review of the decision.

2. Several school principals in the district were phoned to assess their willingness to participate in the study.

3. The Director received permission from the district's Assistant Superintendent and the Chairman of the principals' group to make a presentation about the study at a district's principals' meeting the following week, in order to seek their approval to proceed.

4. A commitment was made by the district's Assistant Superintendent that he would stand by the decision of the principals.

---

*A fictitious name.

As a result of the Director's presentation to the principals, all affirmed their willingness to participate, even though two of the seven principals expressed concern that other school activities might present some scheduling and administrative burdens. Following the meeting, the Director went to the school district offices and relayed the principals' decision to the Assistant Superintendent, who gave his approval to proceed with the study.

Because Evaluation Units are not always so fortunate as to be able to turn a "no" into a "yes", hindsight affords us with the opportunity to have a clearer understanding of what might have been done to avoid the "critical incident" described.

Several thoughts surface:

1. To assume that school districts feel obligated to cooperate with requests from State Department of Education Evaluation Units is a false assumption.

2. Letters which seek to involve a school district but do not specify a procedure for expressing a non-cooperative stance may lead to an unfounded sense of confidence in the Evaluation Unit about how many "for sures" there actually are.

3. Responses to written communications take time and, therefore, a low level of resources should be expended (salaries, printing, secretarial, training, etc.) prior to commitments being solidified between Evaluation Units and school districts.

4. Verbal assurances do not replace the need for written assumrances in situations where key decisions control the outcome of the entire effort.

Problem Case Description #7

The Politics of Evaluation: A Bilingual Case Study

The commitment of the State Legislature and the State Education Agency (SEA) to provide equal educational opportunity to students of limited English language proficiency through bilingual education is reflected in Public Act 78-727. Enacted in September, 1973, PA 78-727 mandated the establishment of transitional bilingual programs in public schools effective July 1, 1976. Prior to this date bilingual programs were conducted by school districts on a voluntary basis. This Act enabled the State Office to provide supplemental financial assistance to local school districts to help them in meeting the costs of their bilingual programs. During that first year of 1976-77 13 million dollars (9,750,000 for Perth* and 3,250,000 for all other school districts) were available.

Transitional bilingual programs are mandated in all attendance centers with 20 or more students of limited English language proficiency of the same language background. Districts with fewer than 20 students as specified may provide bilingual programs on a voluntary basis. Only the "transitional" local efforts are reimbursable by the State Education Agency. Local education agencies can choose to "go beyond" transitional efforts if they are willing to pay the additional cost or seek federal supplementary funding.

During the tenure of the last State Superintendent, the SEA bilingual program administration moved from a one-man staff in the state capital to a staff of fifteen professional and support staff in Perth. The physical move north of 200 miles was sensible programmatically since the vast majority of the students and bilingual programs are in Perth or suburban collar schools. However, this move permitted the bilingual staff to become isolated from many support sections within the agency.

---

*A fictitious name.

The present State Superintendent established the Federal Programs Coordination Council (FPCC) in January, 1976, to develop a clearly defined, centralized intra-agency mechanism for examining the policy impact of federal programs on the total programmatic, administrative, and fiscal operations of the SEA. The Council, composed of the agency's assistant superintendents and several directors, provides a forum for the collective judgments of agency staff responsible for policy formulation on federal matters. Its recommendations are forwarded to the State Superintendent for review and consideration. The FPCC has developed procedures to process federal requests for proposals, agency responses to proposed federal rules and regulations, requests for state endorsements of federal applications and issues.

One of the procedures for quality control that was and still is utilized by the FPCC is a routing slip format. Once initial approval has been given by the Council for a program section to pursue the submission of a proposal or state plan, final approval must be obtained from the Council after a routing slip has been initialed by various sections within the agency. One section that reviews all proposals is the Program Evaluation and Assessment (PE&A) Section.

For a period of one year it seemed that the Bilingual Section had the most serious problems with the Council in obtaining approval of its proposals and especially its staffing plans. After a series of incidents, the directors of the Bilingual and PE&A Sections reached an acceptable arrangement in the hopes of modifying this situation. One of the staff in the PE&A Section in the state capital would have her salary funded out of a combination of federal bilingual plans; this individual would not only perform the required evaluation tasks but would also serve as a the state capital liaison to the Council. In this capacity the evaluation staff member could assist the Bilingual Section staff in drafting proposals and planning.

This resolution had occurred in the late fall of 1976. In the early spring of 1977, staff from the PE&A Section were asked by the Executive Deputy Superintendent to gather and review information for top level management in the agency about the management of the Bilingual Program. Fortunately, that staff member who reviewed the program had a head start in beginning to understand the program. In the following months, the Bilingual Director left the agency and the appropriations bill to fund the entire state program passed the State House by one vote.

Concurrent with this latter series of events, the internal evaluation report produced by the PE&A Section strongly recommended an immediate evaluation to determine whether or not the program was "transitioning" students from the bilingual program to all English classrooms. Yet to be determined was whether these students were really learning English or was this program a way to maintain Spanish language and culture in American schools while simultaneously employing Latino teachers. Not only was there no data available, there was no system to ever produce data; this seemed ironic since without the data system, program personnel would never be able to prove to legislators that the program was really having beneficial effects and was accomplishing its goals. The internal evaluation report recommended that at least an eighteen-month effort be initiated which would obtain preliminary findings and establish a system for future data collection.

However, the final agency decision to have a formal evaluation of the Bilingual Program was delayed beyond the point where the amount could be placed in the annual State Board of Education Budget Request during the late fall of 1977. A decision was made to introduce an amendment to this budget in May 1978 in a State Board committee meeting, then to the full Board and then to the General Assembly. In the meantime, options were discussed internally with regard to the amount of money that would be necessary to conduct such an evaluation and whether it should be done internally by temporary staff or by an external

third party contractor. In either case the decision was made that the Program Evaluation Section would be in charge of the effort and the director of that section would be the project officer. Unfortunately throughout this period of discussion there was no leadership in the program area--applications were being reviewed for the vacant director's position.

The introduction of this fiscal amendment through the State Board went very smoothly. The general Assembly also accepted the amendments as a "quid pro quo" for continued funding of the program. In fact, one of the legislators known for his reservations about the program introduced the amendment for a third party evaluation not to exceed $130,000 for FY 79.

The next step was for the agency to issue an RFP (Request for Proposal) to as large a group of bidders as possible. However, first the questions to be answered in the proposal had to be determined. The Bilingual Program director has been hired by this time; he had several questions. The State Board of Education had questions; so did the SEA Planning Section that is responsible for developing an agency policy position on bilingual education. The General Assembly had questions as well as top level SEA executives. In addition, the Program Evaluation personnel were told that it would be politically prudent to discuss the RFP and the questions with both the Perth Board of Education (since 78 percent of the students are in the Perth program) as well as the Bilingual State Advisory Council, a group established by statute to set directions for bilingual education.

The discussions pointed out the differences among groups and among individuals within groups. Some viewed the effort as evaluation for destruction; others viewed it as evaluation for justification and others as evaluation simply to describe the facts as they existed. Many wanted their own questions included in the RFP at the expense of the questions of the others.

There was every attempt to expedite the development of the RFP and the review process. Although the original recommendation had been for an 18-month project, the section was informed that

practically and politically it would have to be completed during the 1978-79 school year. Ideally, it would be done by mid-December when the State Board was in its budget deliberations for FY 80. Of necessity, some information on the most important questions had to be available by the following May-June debate on bilingual education in the General Assembly.

After the contract was awarded, an Ad Hoc Evaluation Advisory was appointed that contained the following eclectic distribution: Senators and Representatives who are the major declared friends and foes of bilingual education; local district personnel from a rural downstate district, a suburban collar district, and the City of Perth; university experts in bilingual education who had either a qualitative or quantitative background, and the President and Vice-President of the State Bilingual Advisory Council. This group eventually met five times during the eight-month period of the contract. The dialogue among members of this ad hoc advisory panel and between these members and the contractor sensitized everyone about the complexity of the measurement issues.

During the first week of March 1979, when it becomes clear that there would be a successful completion to this contract, the contract project officer held a meeting with the Executive Deputy Superintendent to discuss the agency's approach to responding to the evaluation. It was concluded that the agency response to this evaluation report would be at least as important as the report itself to the General Assembly and their staff. Any hint of defensiveness would be disastrous for program funding and perhaps even its very existence. Instead the agency would use the evaluation report as a springboard to begin making the program changes that everyone knew were needed. The questions did remain though as to how to orchestrate this response.

It was decided that the Bilingual Program director could not chair the response task force. Since there was a possibility that there would be recommendations in the report that, as a program manager he could live with but not as an advocate for his

own bilingual constitutency, the logical choice was someone at the Assistant Superintendent level who had three of his managers closely involved with the project. Therefore, the Assistant Superintendent of Research, Planning and Evaluation was named to chair this task force.

As the deadline for the contractor's draft report approached, there was an attempt to clarify the role of the response task force. Two key players in forming this role were obviously the project officer and the program director. A memo was developed which outlined the sequential five steps necessary for the agency's strategic response to the evaluation.

The steps in this memo proceeded routinely until step four--the presentation of the final draft to the evaluation advisory panel. During that last advisory panel meeting (June 5), several comments were made by SEA program staff that could have been interpreted as attacking the credibility of the contractor, particularly on the collection of the transition data. The time for these questions had been in the technical response to the contractor and not before members of the General Assembly and their staff. These remarks could be interpreted as meaning that there would be "business as usual" and that the evaluation would have little impact.

To avoid this erroneous conclusion, the program director and the project officer were instructed by the Executive Deputy that same day to begin drawing up the SEA response to the evaluation even though the contractor had only delivered a draft report. The focus for the response would be the contractors' recommendation section in the executive summary. The program director returned to Perth and completed the sections concerning program planning and policy; the directors of Research and Statistics, Program Evaluation and Assessment, and Office of Data Management developed responses to the recommendations of a future evaluation plan within a general management information system.

The backdrop for the immediacy in these actions was twofold. First the House had cut 8 million dollars from the SEA request

for Perth and 2.3 million for the downstate request. The Senate needed to know immediately how the SEA was going to react to the external evaluation report so that Senate advocates would have something to point to as they lobbied for the restoration of the funds. House advocates also needed something when the time came for the Senate-House compromise over the appropriations.

Second, there was to be a breakfast meeting for various members of the House and Senate and their staff who were very concerned about bilingual education. This breakfast would occur at 7:00 a.m. on Wednesday, June 13. Originally the subject of this seminar session was to be collective bargaining. However, the governor's office requested that a substitute issue be found and bilingual education became the topic. A panel of three members of the General Assembly would react to the evaluation findings. It became imperative that they also have the simultaneous opportunity to react to the SEA's proposed recommendations or plans to implement the contractor's recommendations.

On June 6 and 7 the first drafts of the SEA reaction papers were written. The Project Officer merged the sections on June 7 and sent them to the Executive Deputy for review. On June 8 the comments for revision from the Executive Deputy and Superintendent were incorporated into a final draft which was approved on Monday morning, June 11. Copies of this reaction were then mailed to every member of the General Assembly except those panel members; their copies were hand delivered.

The strategy worked. The breakfast participants were impressed with the speed and quality of the response. The final appropriation was 4 million dollars for downstate programs (an increase of 400,000 from the previous year) and 12.6 million for Perth, an increase of 1.6 million).

Problem Case Description #8
Quality Control

The Education Bureau is responsible for the collection of data on student achievement for students who are in categorically aided programs. Individual data are collected in the areas of reading, math, writing, and bilingual programs. Reports are then prepared by the Bureau in compliance with state and federal mandates.

The Bureau is faced with the problem of quality control in ensuring that the data collected are accurate and usable on the local, state, and federal levels. Before the data reach the state's Education Bureau, the data may have been handled by many persons. For example, the reporting forms may have been filled out by several different teachers, submitted to the principal, forwarded to the chief school officer, and then forwarded to the superintendent of the Region who, in turn, submits several district's reports to a Regional Computer Center where the data are entered on a magnetic tape to be sent to the State Education Department's Bureau of Evaluation. Problems arise when the data received by the Bureau are not machine processable. Since the Bureau uses a computer, the data received must be in an acceptable format. School districts must complete several forms on each student and the number of forms and type of information requested vary in accordance with the type program a student was enrolled in and the evaluation design the district chose for that program. In the past, the Evaluation Bureau has spent many months screening the data and correcting such gross errors as the use of the wrong district code or an incorrect or missing card identified. These errors are the tip of the iceberg. Many errors, which the Bureau cannot correct without contacting the districts for the correct information, skew the results. The Bureau has developed a list of errors which repeatedly occur and has informed school districts that they will be responsible for the correction of their data errors this year.

The following is a partial list of types of errors which will be detected by the computer program:

a. Improper population code, e.g., nonpublic school pupil assigned to a public school building.

b. Improper component code, e.g., impossible code number.

c. Improper test code used for both norm- and criterion-referenced tests.

d. Improper test level, e.g., 2nd grade student given high school level test.

e. Improper month for pre- or posttest, i.e., out of the norming period.

f. Table missing.

g. Raw test score missing for pre or posttest.

h. Normal curve equivalent or Percentile missing for pre or posttest.

i. Birthdate missing for pupil in an ungraded class.

j. Duplicate data on a pupil.

k. Improper sub test code, e.g., vocabulary code used with mathematics test.

Errors are classified into three types. Type 1 errors are critical errors and they must be corrected for any analysis of the data. Type 2 errors are considered substantial and they must be corrected for more meaningful analyses. Type 3 errors are classified as "information". Generally, they cannot be corrected; however, they may alert recipients to the need for a modification in testing procedures or a change in the evaluation design in the next school year. A Type 1 Error would be a missing district code. A Type 2 Error would be when a math component is listed for a pupil, but the number of contact hours in Math for that pupil is missing, and a Type 3 Error would indicate that the test administered may be too difficult for the student.

The Bureau collects information on approximately one million pupils. The task of data collection is enormous for both the Evaluation Bureau and the local education agencies. Assistance is needed in the area of quality control and the development of a system whereby the data are screened along various check points will help to ensure the data received by the Bureau are as accurate as possible. The error correction procedure is an attempt to begin this process. Additional elements need to be added to the error correction procedure to complete the quality control system. The implementation of the error correction procedure and other quality control elements are essential to the rapid processing of data that will make possible the timely return of information to local education agencies, the state education agency, and the federal government. Without timely return, the utility of evaluation for decision making purposes is lost and the evaluation simply meets reporting requirements.

Problem Case Description #9
Basic Skills Evaluation

Introduction

At the end of the last legislative session, a bill was passed which dealt with Basic Skills. One of the requirements of the legislation was to evaluate the state Basic Skills program. Funding for the program was for staff, staff expenses, and in-service costs. The evaluation budget was zero.

This bill was negotiated at the "last minute" primarily by a few legislators and with minimal consultation with the State Department. In a sense it was a very small token for the Governor who had wanted another very large education bill.

General Background

There were many logistical problems in getting this program organized and off the ground. In the Department there were two schools of philosophy. One group thought basic skills were as defined in the legislation, i.e., abilities to listen, speak, read, write and compute. Another group, composed of subject area consultants, felt basic skills were an integral part of every subject, i.e., science, and thus they should be involved in the statewide planning. The issue was finally resolved and the former philosophy prevailed.

A Department task force was developed with people from each Division of the Department to assure across-Department coordination and to guide program direction. A Basic Skills supervisor was hired along with 11 Basic Skills specialists. These specialists were assigned to work in regions of the state in intermediate unit offices.

However, the regional specialist reported to Department even though they were housed in and worked with staff of regional units. This was "upsetting" to many directors of regional units who felt if a person were housed in his office he/she should have

some administrative responsibility. Other state and federal programs had provided staff to intermediate units and all were under the direct administrative control of the regional unit director.

The operational model for the Basic Skills program is the following:

1. All Basic Skills specialists are trained simultaneously.

2. After training, the specialists will then train staff from districts who wish to participate in the program.

3. After training, the local school persons trained will then implement a basic skills program that reflects the Basic skills Standards of Excellence.

Another issue was the reluctance to make the standards too specific. There was a feeling that perhaps if the standards were kept more general, accountability might be easier.

Assignment of Evaluation Responsibility

No evaluation funds were included in the legislation. Staff of the Evaluation Unit in the Department was assigned the evaluation responsibility as an additional task to its current activities.

Program Implementation

The Basic Skills program was modeled after a previously highly successful program "Right to Read". Each regional person would train persons from local school districts to be the leader and coordinator of Basic Skills activities in their districts. Persons from local schools would be trained in both content and process. From December to April, regional specialists would explain the Basic Skills program to local schools and encourage them to participate. From April to June there would be training of persons from the volunteer schools. Training also would take

place again after the summer. The local schools would then begin to implement the Basic Skills program.

The Basic Skills program had a series of 18 statements, some measurable and some not, called "Standards of Excellence". These were the standards to be used which defined what an ideal comprehensive Basic Skills program would include.

Preliminary Evaluation Tasks

Problem 1. To determine what the legislators had in mind as evaluation outcomes of the program.

Action. Prepared a proposal that suggested that the principal authors of the legislation be consulted and shown some alternative questions which might be reasonably answered in the next year (since a report had to be made to the legislature by the following January, six months before the funding and program activity would be completed and funding would end). This would attempt to make decision makers more involved in the evaluation process and possibly encourage them to utilize evaluation results for decision making. It would also clarify what kind of information was needed at their level in contrast to the kind of data that is needed for program management evaluation. This procedure was rejected by Department decision makers. As a result, an evaluation plan was prepared which may or may not be of interest to legislators who will determine the continuation of the project.

One major issue well could be that only very preliminary information will be made available by January and that the legislators are expecting outcome data.

Problem 2. The Standards of Excellence were generally vague and not measurable.

Action. After a series of meetings with the Department Committee and Basic Skills specialists, measurable criteria were established for each Standard. It was important that these be measurable so that data could be obtained from all participating districts as base line data.

One major issue in this area is that since there would be no new basic skills activities in school districts until four months before the preliminary report to the legislature was received, very few new activities probably would be implemented. It is critical that the report clearly explain the meaning of base line data.

Problem 3. Can any new basic skills activities in school districts be directly attributed to the new Basic Skills program.

Action. There are other state and federal programs relating to basic skills. There is no way to clearly establish a direct relationship between this new program and improvements in basic skills programs.

Problem 4. Was the training of local school persons effective?

Action. All workshops needed to be evaluated to determine if the participants were learning the skills and processes to implement a coordinated integrated basic skills program in their district. A workshop evaluation questionnaire was developed (see Exhibit 3).

A major issue was the concern expressed by the Basic Skills staff, many of whom had little experience in putting on workshops or in having their "performance" evaluated. An administrative decision was made, however, that the evaluation would be done.

Problem 5. How to measure the impact of the Basic Skills program on improved basic skill test scores.

Action. It would be impossible to establish a measurable relationship between a regional training program and student achievement. Changes in basic skills scores will be monitored through the State Assessment Program, over at least five years after the inception and hopeful continuation of the Basic Skills program.

Problem 6. If no hard data are available how can it be determined if the Basic Skills program appears to be making positive changes in school districts?

Action. A process evaluation plan was developed. An annual survey would be done on a yearly basis; the responses would be based upon the professional judgments of the respondents: lay and advisory council persons, school board persons, school superintendents, and school staff. Their feelings would be the principal data which could be shared with decision makers as an indication of program success.

A major issue always is the use of judgment and opinion data. The precedent for this process had been established and accepted in the former Right to Read program. Because of the similarity of the models of the two programs (Basic Skills and Right to Read) judgment data should be accepted. However, if the current positive climate in the state toward education changes, the impact of judgment data could be reduced. In programs such as this where the outcomes are not clearly defined and in which the content areas cuts across many other ongoing educational activities, it is difficult to gather hard data.

Problem 7. Funding for evaluation activities.

Action. Miscellaneous evaluation expenses needed to be covered by other program funds. Substantive evaluation is, therefore, very much constrained.

Evaluation Summary

An evaluation needed to be done. For various reasons the evaluation staff was unable to meet with the legislative decision makers. This was an example of the role of politics in evaluation. Program evaluation at a state level often is a long way away from the "textbook" evaluation procedures taught in colleges and universities. Knowing that the evaluation that was to be done was going to be less effective than it could have been perhaps raises the issue of the professional integrity of the evaluator versus a need for survival in his/her job.

Will the proposed evaluation activities have any influence on the final decision for the continued funding of the Basic Skills

program? Probably some, but not a major factor. Only when evaluation is considered when legislation is developed can evaluation results have a chance of being a major factor in continued fundings of programs.

Problem Case Description No. 10

Monitoring - A Threat to Evaluation?

Things started to go wrong on Jack's third visit to the program funded by his agency. He felt he had lost all rapport and cooperation with the program people he was to help. Even visits to other funded programs were not well received. It was as if word was out to beware of Jack.

During his first visit, Jack thought he was accepted as an evaluator. He had worked with the program director to show that evaluation, with its myriad of definitions dependent upon the different schools of thought, may be viewed as the assessment of the value or worth of a program or activity. He felt that evaluation could be seen as a tertiary relation, (x, y, z), where "x", as an evaluator, acts with a set of data "y" to determine whether a standard "z" is met. He knew that decisions based upon evaluation results should be made by the program director or administrators.

Jack was also pleased with his second visit as a monitor. He explained to the program director and others that monitoring may be thought of as a process for ascertaining whether a program or activity was within various rules, regulations, minimum standards, or agreed upon terms. Monitoring also may be seen as a tertiary relation, (x, y, z), where "x", as a monitor, acts with a set of data "y" to determine whether a prescribed precept "z" is met. Jack knew that decisions based upon monitoring results were made by the funding authorities and policy makers.

Jack was well aware of the strain during the third visit when the program director asked, "For what reason are you here? Are you here to help us with our program or are you here to be a watchdog?" In the eyes of the program director, Jack, as an evaluator, suddenly became a monitor.

In the case of Jack, the roles of an evaluator and a monitor became confusing. At one point, he was performing activities for purposes of evaluating the program, while at another time, he was

collecting information or observing for monitoring purposes. Was this sensed lack of rapport due to his agency having the same unit or persons responsible for both evaluation or monitoring? Although Jack probably had no problem with the role change since the focus of a monitor and evaluator are different, the funding recipient had difficulty in interpreting the intent of the visits. In effect, a relazed atmosphere necessary for communicating clear information to an evaluator was lost.

Jane, an evaluator in another agency, also ran into problems with her visits and communications with funded project personnel. She was attempting to assist in setting up evaluation designs and collecting formative information when asked by the project personnel, "How are you going to use the project information?" Jane believed that she had clarified earlier how data were to be used in an evaluation process and was surprised at the further line of questioning, "Will the information be used to assist in the improvement of the program or will the information be used to cut or reduce funding?"

With further investigation, Jane found that the project received a monitoring instrument by her agency. The project personnel felt that many of the questions in the instrument were too similar to those being solicited in the evaluation. The purpose or use of the information became so unclear that the project personnel were afraid that some information may be misinterpreted and used against them. They were fearful that the program could be in jeopardy if they answered questions in a frank manner, especially when some adjustment were presently being made to alleviate identified problems.

As in Jack's situation, the atmoshpere necessary for communicating clear information was lost. Jane had to spend a large amount of time assuring that information she collected would be used at the agency level for assessing the value or worth of aprogram and for providing feedback to assist in the development of the program. She further assured that her data would not be used for monitorning compliance to various rules or regulations.

Problems between evaluation and monitoring can occur whenever various corresponding components of the two tertiary relations are equal or elements within the components intersect. Jack's case illustrated a situation in which the first components were equal. In this case, the evaluator and the monitor were the same person. Jane's case of having collected some evaluation data similar to that collected on a monitoring instrument exemplifies a situation whenever elements within the components of the relations intersect (elements of the second component intersect). This situation can also occur whenever the same unit within an agency is responsible for both the evaluation and the monitoring of programs (elements of the first components intersect). Although it is theoretically possible for other component equality or intersection of elements, this author has not observed such instances. This absence may be a good sign as occurrences of other instances would indicate serious problems with understanding the differences between evaluation and monitoring.

An ideal situation would be the case where no equality or intersection occurs. However, with organizational structures which place evaluation at a low priority, with the limited knowledge base of various persons about evaluation and monitoring, and in a time of staff reduction, it is very unlikely that the ideal situation will occur in many funding agencies. Thus, to lessen problems between evaluation and monitoring, strategies need to be developed for situations where corresponding components within the tertiary relations are equal or where elements within the components of the relations intersect.

Strategies should be planned for addressing problems at both the funding agency and the funded agency levels. Answers need to be found for such problematic questions as: How does one assure to funded project personnel that evaluation results will be used to judge the worth of quality of programs? What does it mean to judge the worth or quality of a program when the question of

continued funding is in the minds of funded project personnel? What can be done to better delineate the roles of evaluators and monitors? What are some ways a funding agency can assist staff when the evaluator is also the monitor or is in the same program unit as is the monitor? What are the major differences in evaluation data and monitoring data, and how can these differences be best communicated to those program policy makers or decision makers who are not evaluators?

Although some answers to problems may seem quite evident to ones who are knowledgeable in evaluation and who, for example, can distinguish between evaluation and monitoring, one must keep in mind that the evaluator is but one actor out of many who are involved with programs. As in Jack's case, the problem was not necessarily with Jack having to differentiate between the role of an evaluator and a monitor, but rather, with the funded project staff feeling that a conflict of roles existed. Even though the funding agency may recognize a difference between an evaluator and a monitor, a solution to the problem by having different evaluators and monitors may not be that simple when funds are not available to hire additional qualified staff. Jane knew how to apply her data against evaluation standards, but the project staff was not convinced. Perhaps they felt her superios may misuse the evaluation results and make regulatory decisions or funding decisions.

Monitoring--is it a threat to evaluation? Perhaps not, if possible problems are anticipated and strategies developed to address these problems.

Problem Case Description No. 11

Betwixt and Between

Background

For State Offices of Education, accepting federal funds often carries with it a responsiblity to gather evaluative data. The data then are reported to the appropriate federal agency.

Recently, however, in one area--Title I 89-10--federal meta-evaluation studies concluded that evaluation reports received from states were of doubtful quality and that no aggregable national data could be compiled because of the diversity of the reports. The consequence was that in Title I 89-10 and other programs, federal evaluation guidelines now include the prescription of acceptable evaluation models. It is argued that implementation of such evaluation models as these will provide information which will allow sound generalizations to be made at the national level. This may be true, but in the implementation of the models a situation is created where the State Office evaluator is caught betwixt and between providing a service basically for the benefit of a federal agency, providing evaluation services useful to State Office decision makers, or providing services to Local Educational Agencies which has their benefit as a major concern.

Nature of the Problem

When providing evaluation for federal programs the major evaluation functions of the State Office of Education are often three-fold. One function is to provide service to Local Educational Agencies by providing assistance which would be compatible with local needs and which would add incrementally to the current lo l level of sophistication. The second function, serving State Office administrators, includes gathering data regarding compliance with federal law, gathering information which would forecast potential troublesome areas, and gathering

data which would complement grant writing efforts. The third function is to serve the "Feds". This includes monitoring LEA evaluation efforts, making the required annual reports, serving as an intermediary between the Feds and the LEA; and providing technical assistance in the implementation of the federal evaluation models at the local level.

Of these three sets of functions, the primary concern of the State Office is to serve federal interests. That primacy derives from the power of the purse. Because federal monitoring is a possibility, the emphasis within the State Office is to serve federal functions at the expense of any other interests.

At this point a variety of complications should be mentioned. Most of the complications reflect the limited amount of administrative funds allocated to the state.

The constraint on funding has a variety of ramifications. federal funding is administered in this State Office by program staff who are separate from Evaluation staff. Program staff typically argue that allocation of administrative funds needs to be directly supportive of local program impact upon children. At that point, program staff point out that State Office evaluators primarily function as a service for federal interests. Evaluation, then, is assigned a low priority by Program staff.

Limited funding assures that only one or two staff are allocated to evaluative purposes at the state level. In addition to a limited number of personnel for evaluation, travel and additional resources are limited.

The lack of staff provides a severe constraint on the scope of work possible. With federal interests primary, state evaluators often put other interests aside completely. In the eyes of the program administrator this confirms the low priority assigned to evaluation. For the LEA, it virtually assures their frustration when anything but federal model evaluation is broached with State Office evaluators. For LEA staff uninitiated to the scope of evaluation, it creates the impression that experimentalist and measurement oriented outcome evaluation is

what evaluation is all about. For small schools or schools lacking technical sophistication, the models are viewed as burdensome and without potential transferability to other evaluative purposes.

Parameters for Resolution

The quick and easy dismissal of the "problem" as being remediated by reapportionment of State Office funds to provide for additional evaluation staff is not likely to be possible. As was indicated, evaluation is perceived to be a low priority for the persons administering the funds. Something must be done first by the State Office evaluators which will demonstrate the benefit of evaluation services.

At the State Office level the demonstration should relate to such aspects of expected functions as forecasting potential trouble areas. This must be done with limited travel, limited staff while at the same time the staff fulfills obligations to assist in the implementation of federal evaluation models.

In support of LEA evaluation efforts, considering many of the same points listed in the previous paragraph, what can the State Office evaluator do? Here the undertakings must be flexible enough to adapt to a variety of LEA setings--including such features as size, technical sophistication, available local staff, and restricted finances.

Another quick dismissal of the "problem" might be attempted by expanding the funding and functions of the Technical Assistance Center idea as implemented in Title I 89-10. But TAC's are constrained by their basic function of supporting the implementation of the federal models. Evaluation needs go beyond those models. The evaluation needs at the State Office level also would not be typically within the scope of interest of the TAC's. Finally, the TAC's often represent a high technology approach to the solution of problems. High technology solutions are beyond many of those in need of evaluative assistance.

It may nearly be an impossible task to devise evaluative tactics which can be circumscribed within such bounds. I believe, however, that much in these parameters delineate a great need for evaluation. We, as evaluators, must have a kit bag which includes evaluation approaches or tactics which can be implemented by a few persons, in a short time, with limited time and funds; and in addition, many times to be implemented by persons or relatively modest evaluative sophistication.

In our day with our fascination with high technology, it is easy to forget the multitude of evaluative needs which will not be thus served. The forgotten needs will be those where money and power are lacking--yet they are not insignificant. It seems to me that it is there where much will be done to improve the condition of education.

Problem Case Description No. 12
Maximizing Assessment Results Utilization

Probably one of the most vexing problems existing for State Departments of Education are attempts to make assessment results meaningful at multiple levels of the educational hierarchy. Teachers want to use assessment results for student diagnostic and perscriptive purposes, while building administrators look to assessment results for within building program and curriculum evaluation. District personnel perceive assessment results as a way of evaluating programs and curriculum between buildings, districtwide. At the state level, assessment results are viewed in a broader context. Typical state utilization of assessment data are for evaluation of students, programs and districts across the state. Often domain deficiency within content areas are explored. The main utility of assessment results, however, at the state level is for supporting evidence for additional funding by legislatures or policy issues by State Boards of Education. Most statewide assessment programs focus on the latter uses of results, often overlooking teachers', building administrators' and districts' needs for quick accurate assessment results for evaluation of students, curriculum or programs at the local level.

Paramount among the problems of multiple use of assessment results is the scoring of answer sheets. Scoring systems seldom offer feedback to local school districts or teachers, such that this information can be used for local decision making. Usually student responses to questions on statewide assessments are obtained under some machine readable format, optical scan answer sheets being the most popular. The answers are collected by the teacher or test administrators, send to the building level for aggregation of all classrooms assessed in that building, and then sent to the school district for further aggregation of districtwide assessed classrooms and finally to the state for scoring of the answer sheets. In many cases states send the

answer sheets to a scoring service for processing. The results of the statewide assessment flows in reverse order in reporting the results back to the local school districts. This scoring and reporting process can take from several weeks to several months. Therefore, if the assessment results are reported in such a format which can be useful on the local level (often they are not), they do not filter down to the appropriate level in a timely fashion. The irony of the statewide assessment scoring process is that those that have the need for immediate test results, i.e., making student decisions, often are the last to receive the results, while those that are more interested in the overall picture of statewide assessment and are unable to act on the results until next legislature or State Board convenes are the first to receive the results.

The assessment reports which are generated from the answer sheets require different information for each of the levels of utility. There is the need for very detailed and discrete information on the reports which are returned to the teacher in order for her/him to make judgments about students. These reports must indicate item by item how each student performed so that the teacher may make judgments about students and design classroom activities which will facilitate student achievement. This report is very different from, say, the building level report, district report, or even the state reports which are more concerned with averaging of scores across domains, reporting items for item analysis or reporting of gross scores across programs, buildings or districts. Usually as one moves up the educational hierarchy, assessment information is needed in a decreasing order of detail.

Finally, there is always the problems of confidentiality as the reporting moves away from the classroom. If the assessment data is to be collected such that teachers may use the results for classroom purposes students must be identified. At the state level there is little if any need to identify individual students, however, certain student characteristics may be very

important (sex, race, handicapping conditions of students). While reports that do not identify students, buildings or classrooms may have little local utility, these identifiers are usually not needed for any statewide assessment. State or federal confidentiality constraints or the degree of confidentiality required of individual students may impede assessment efforts at the state level in the future and therefore limit the utility of statewide assessment reporting.

The utility of assessment results at the various educational levels needs to be addressed by those doing the assessing, if for no other reason than to offer some guarantee that assessment results have a degree of validity. It is not enough anymore to expect local school districts, schools, and teachers to enthusiastically administer a statewide assessment because "they" need the data for "some" reason. Only when those administering the test instruments to students can plainly see some direct benefit can we hope for compliance to standardized administrative methodology. The validity of statewide test results becomes particularly troublesome when teachers must administer a second instrument to obtain the data that will be useful for their decision making.

A suggested solution to this problem is to have assessments initiated at the local school district, school or classroom level and then the results of each of these aggregated as they are sent forward to the state. The problem with this approach is the lack of coordination in order to obtain comparable data to aggregate. There is a need for agreement on what is to be assessed, how and when the assessment is to take place. Different testing times, nonstandardized assessment procedures and attempts in aggregation of assessment results where instruments are not comparable are major problems.

At first blush the above solution seems to have little or no merit. However, if examined for technical feasibility as well as practicality, it is quickly recognized that the current state of the art of scoring and the rethinking of out moded ideas can make

this solution a viable direction in which to pursue. The solution as posed, rests on theoretical concepts which we have accepted over the years due to past technical limitations. Mainly, the scoring of student tests must take place outside the school, or school district. This was due to the fact that machinery to score student tests was very expensive. With the new computer technology, it is within the grasp of most schools and school districts to do their own scoring and report generating. With assistance from state and federal governments assessment support materials could be purchased (computer hardware and software) that would produce a new type of assessment administration that would furnish appropriate assessment results at every level of the educational hierarchy. This method would require that the student assessment data be scored and reported at either the school or school district level and then data sent forward in a machine readable format, aggregated as it moves toward the state level. The instrumentation could be a joint effort by state and school districts of identifying areas to be assessed and other specifics. This method of assessment would maximize the utilization of results. Those that need the results earliest and in the greatest detail would have first call on the results, while those that need less detail but a broader data base would also obtain results in a timely fashion.

Problem Case Description No. 13
Coleman Assessment Program

Background

There has been a legislatively mandated state-testing program in Coleman since 1962-63. The Coleman Assessment Program (CAP) as it is currently known, tests all public school students annually in grades 1, 3, 6, and 12.

The first grade test is an entry level test administered at the beginning of school to measure the readiness skills which pupils bring with them to school. In the other grades state-developed tests are administered on a matrix sampling basis to monitor the achievement in reading, language, and mathematics. No individual student scores or classroom reports are produced. The smallest unit of analysis is the school. School data are then aggregated to produce district reports and an annual report of statewide achievement.

A state director and six consultants administer the program at the state level. Contracts are let for printing, distribution, collectiohn, scoring, analysis, and preparation of reports which are mailed to each of the 1,044 Coleman school districts. Regional workshops are held each fall to acquaint LEA personnel with the content and interpretation of the reports. All assessment information is available to the public after it has been presented to the State Board of Education in November following the school year in question.

School and district reports include the raw scores for the current and prior school years. Look-up tables provide norms in the form of state percentile ranks. Background factors, such as socioeconomic indexes and English language fluency, are reported and used in multiple regression analysis to produce a predicted score. Reports also include student score distributions and subscores by skill area (e.g., capitalization and punctuation) and by subpopulation (e.g., boys vs. girls).

There are many issues, difficulties, dilemmas, which an assessment program faces. Few of those can be resolved to the satisfaction of all those affected by the program, but every one must be addressed. Each of the following paragraphs touches upon an issue with which CAP staff must deal. While the listing is not even exhaustive, it can nonetheless mislead the reader into concluding that CAP has only problems. CAP has been notably successful. However the purpose of this paper is not to look at successes but rather to focus on problems. Hence the following focuses on the empty portion of the glass, not the full portion.

Problems

1. The dilemma of reconciling local autonomy in curriculum with a state program which assumes, or seeks, commonality of teaching objectives. Coleman does not have a prescribed state curriculum. Thus it is difficult to achieve consensus among educators as to what should be taught and thereby should be assessed.

2. A large-scale assessment program must rely almost solely on paper and pencil, machine-scorable, multiple-choice tests. While many teaching objectives lend themselves to such measures, many other skills--particularly in the minds of many educators--can be less directly assess in such a mode.

3. Shifting legislation and financial support make long-term planning tentative at best. Although CAP is legislatively mandated, those mandates have changed many times in 18 years--which grade levels are assessed, what content areas are tested, switch from requiring IQ tests to prohibiting them. Likewise budget allocations are unknown until the last minute, particularly in recent years with state initiatives such as Propositions 13 and 9.

4. How to motivate students to put forth their best effort. Because individual students have no stake in the results, there is little incentive for them to do well on the tests, particularly so with high school seniors.

5. How to distinguish between true instruction and coaching or cheating. While the state wants to encourage schools to teach the objectives of the test and particularly to remediate skills previously found to be weak, sometimes these efforts could be viewed as teaching the test items rather than teaching the test objectives.

6. How to meet the conflicting goals of comparative information and instructionally relevant information. Some audiences, such as legislators and the press, demand normative (ranking) information for schools and districts. Those schools which rank relatively low become hostile and therefore aren't willing to look beyond the negative ranking to discover in what skills their students are deficient.

7. How to keep assessment reports simple and usable without sacrificing completeness and accuracy. The recipients and users of the reports differ widely in their technical, statistical sophistication. It is difficult to reconcile the dilemma of keeping a report simple enough to be usable by the classroom teacher without offending the research staff of larger school districts who would label such a report misleading and simplistic.

8. How to encourage use of assessment information at the local school level. The complexity of the current reports (cited above in number 7) is only one factor inhibiting usage. Other factors are even more difficult to overcome--e.g., antipathy of teachers to testing, the opposition to what is viewed as state control or state interference, feeling that teachers themselves are being evaluated, frustration at repeatedly unsuccessful efforts to improve achievement in schools serving disadvantaged students.

9. Whether to keep tests constant to encourage longitudinal analyses or change tests to improve them. LEA Personnel continually request the conflicting need to improve the tests yet keep them the same. How responsive should the state be to recommendations or suggestions? it is difficult to ascertain from isolated comments or committee recommendations how representative those thoughts are.

10. How to deal with scores of students in special programs such as special education or bilingual. The advent of mainstreaming for special education students and the increasing number of students whose English language fluency is limited have magnified the dilemma of whether to include their scores in those for the school.

11. Relation between assessment and minimum competency testing. Thus far these two programs have been viewed as two distinct programs; however, the layering on of additional tests, which results in duplication and triplication of testing of students (who are in ESEA Title I) is all too obvious to LEAs and a method of consolidating testing will be in demand to save money and instructional time.

12. Education legislation usually has no sanctions for non-observance. The CAP, as is the case with most education legislation, seldom has penalties for non-compliance. (The penalties, if they exist, are seldom invoked.) The program's success relies upon the acceptance and cooperation of LEAs. If a LEA tests, say, only 75 percent of its students, the state has no viable remedy or method of encouragement or coercion to force them to test something closer to 100 percent.